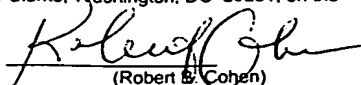


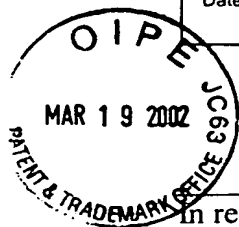
I hereby certify that this correspondence is being deposited with the U.S. Postal Service with sufficient postage as First Class Mail, in an envelope addressed to: Commissioner for Patents, Washington, DC 20231, on the date shown below.

Dated: March 11, 2002 Signature: 

(Robert B. Cohen)

#4

Docket No.: SCEITO 3.0-096  
(PATENT)



**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Patent Application of:  
Terasawa et al.

Application No.: 10/013,057

Group Art Unit: 2641

Filed: October 30, 2001

Examiner: Not Yet Assigned

For: ENTERTAINMENT APPARATUS AND  
METHOD FOR REFLECTING INPUT VOICE  
IN OPERATION OF CHARACTER

Commissioner for Patents  
Washington, DC 20231

**CLAIM FOR PRIORITY AND SUBMISSION OF DOCUMENTS**

Dear Sir:

Applicant hereby claims priority under 35 U.S.C. 119 based on the following prior foreign application filed in the following foreign country on the date indicated:

<u>Country</u>	<u>Application No.</u>	<u>Date</u>
Japan	P2000-330800	October 30, 2000

In support of this claim, a certified copy of the original foreign application is filed herewith.

Dated: March 11, 2002

Respectfully submitted,

By 

Robert B. Cohen

Registration No.: 32,768

LERNER, DAVID, LITTENBERG,  
KRUMHOLZ & MENTLIK, LLP  
600 South Avenue West  
Westfield, New Jersey 07090  
(908) 654-5000  
Attorneys for Applicant

**COPY OF PAPERS  
ORIGINALLY FILED**



日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日  
Date of Application:

2000年10月30日

出 願 番 号  
Application Number:

特願2000-330800

出 願 人  
Applicant(s):

株式会社ソニー・コンピュータエンタテインメント

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年 6月 5日

特許庁長官  
Commissioner,  
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3052597

【書類名】 特許願

【整理番号】 SCEI00042

【提出日】 平成12年10月30日

【あて先】 特許庁長官殿

【国際特許分類】 A63F

【発明の名称】 入力音声をキャラクタの動作に反映させるエンタテインメント装置、方法および記憶媒体

【請求項の数】 21

【発明者】

【住所又は居所】 東京都港区赤坂7丁目1番1号 株式会社ソニー・コンピュータエンタテインメント内

【氏名】 寺澤 健治

【発明者】

【住所又は居所】 東京都港区赤坂7丁目1番1号 株式会社ソニー・コンピュータエンタテインメント内

【氏名】 白須 昇

【特許出願人】

【識別番号】 395015319

【氏名又は名称】 株式会社 ソニー・コンピュータエンタテインメント

【代理人】

【識別番号】 100084032

【弁理士】

【氏名又は名称】 三品 岩男

【電話番号】 045(316)3711

【選任した代理人】

【識別番号】 100087170

【弁理士】

【氏名又は名称】 富田 和子

【電話番号】 045(316)3711

【手数料の表示】

【予納台帳番号】 011992

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9912211

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 入力音声をキャラクタの動作に反映させるエンタテインメント装置、方法および記憶媒体

【特許請求の範囲】

【請求項 1】

プレイヤーからの音声入力を受け付ける音声入力装置を利用可能なエンタテインメント装置であって、

ゲームキャラクタの動作を制御するキャラクタ制御手段と、

上記音声入力装置を介して受け付けたプレイヤーの音声から、相対音程に関する情報を抽出する音程抽出手段と、

前記音声入力装置を介して受け付けたプレイヤーの音声から、音量に関する情報を抽出する音量抽出手段とを備え、

前記キャラクタ制御手段は、前記抽出した相対音程に関する情報と音量に関する情報とを評価し、その評価結果に応じた動作をキャラクタに行わせることを特徴とするエンタテインメント装置。

【請求項 2】

請求項 1 記載のエンタテインメント装置であって、

プレイヤーが入力すべき音声の内容を出力するガイド表示手段をさらに備えることを特徴とするエンタテインメント装置。

【請求項 3】

請求項 2 記載のエンタテインメント装置であって、

プレイヤーが入力すべき音声に対して、相対音程および音量についての評価基準となる音声データを記憶する基準音声データ記憶手段を備え、

前記キャラクタ制御手段は、前記抽出された相対音程に関する情報および音量に関する情報と、前記評価基準となる音声データとを周期的に比較し、その比較結果に基づいてキャラクタの動作内容を決定することを特徴とするエンタテインメント装置。

【請求項 4】

請求項 2 記載のエンタテインメント装置であって、

プレイヤーが入力すべき音声の表現態様を出力する表現態様表示手段をさらに備えることを特徴とするエンタテインメント装置。

【請求項 5】

請求項 3 記載のエンタテインメント装置であって、

前記キャラクターの動作は、あらかじめ用意された画像データを再生することにより表され、

前記キャラクター制御手段は、前記プレイヤーが入力すべき音声の内容を出力するタイミングと、プレイヤーが音声の入力を開始するタイミングとの差に基いて、前記画像データの再生速度を変化させることを特徴とするエンタテインメント装置。

【請求項 6】

請求項 3 記載のエンタテインメント装置であって、

前記キャラクター制御手段は、前記抽出された相対音程に関する情報と、前記評価基準となる相対音程に関する音声データとを比較した結果、抽出された相対音程が評価基準となる相対音程より高いほど、キャラクターの表情を大げさにし、抽出された相対音程が評価基準となる相対音程より低いほど、キャラクターの表情をひかえめにすることを特徴とするエンタテインメント装置。

【請求項 7】

請求項 3 記載のエンタテインメント装置であって、

前記キャラクター制御手段は、前記抽出された音量に関する情報と、前記評価基準となる音量に関する音声データとを比較した結果、抽出された音量が評価基準となる音量より大きいほど、キャラクターの振る舞いを大げさにし、抽出された音量が評価基準となる音量より小さいほど、キャラクターの振る舞いをひかえめにすることを特徴とするエンタテインメント装置。

【請求項 8】

エンタテインメント装置で実行するゲームにおけるキャラクターの動作を制御する方法であって、

プレイヤーの音声入力を受け付けると、その音声データから相対音程に関する情報と音量に関する情報とを抽出し、

前記抽出した相対音程に関する情報と音量に関する情報とに基いて、キャラク

タの動作を変化させることを特徴とするキャラクタの動作を制御する方法。

【請求項 9】

請求項 8 記載のキャラクタの動作を制御する方法であって、  
 プレイヤの音声入力を受け付けに先立ち、  
 プレイヤが入力すべき音声の内容を表示することを特徴とするキャラクタの動作を制御する方法。

【請求項 10】

請求項 9 記載のキャラクタの動作を制御する方法であって、  
 前記キャラクタの動作の変化は、  
 前記抽出された相対音程に関する情報および音量に関する情報と、あらかじめ用意された相対音程および音量についての評価基準となる音声データとを周期的に比較し、その比較結果に基づいて決定されることを特徴とするキャラクタの動作を制御する方法。

【請求項 11】

請求項 9 記載のキャラクタの動作を制御する方法であって、  
 プレイヤの音声入力を受け付けに先立ち、前記プレイヤが入力すべき音声の内容とともに、プレイヤが入力すべき音声の表現態様を表示することを特徴とするキャラクタの動作を制御する方法。

【請求項 12】

請求項 10 記載のキャラクタの動作を制御する方法であって、  
 前記キャラクタの動作は、あらかじめ用意された画像データを再生することにより表されるものであり、  
 前記プレイヤが入力すべき音声の内容を出力するタイミングと、プレイヤが音声の入力を開始するタイミングとの差に基づいて、前記画像データの再生速度を変化させることを特徴とするキャラクタの動作を制御する方法。

【請求項 13】

請求項 10 記載のキャラクタの動作を制御する方法であって、  
 前記抽出された相対音程に関する情報と、前記評価基準となる相対音程に関する音声データとの比較の結果、抽出された相対音程が評価基準となる相対音程よ

り高いほど、キャラクタの表情を大げさにし、抽出された相対音程が評価基準となる相対音程より低いほど、キャラクタの表情をひかえめにすることを特徴とするキャラクタの動作を制御する方法。

【請求項 1 4】

請求項 1 0 記載のキャラクタの動作を制御する方法であって、

前記抽出された音量に関する情報と、前記評価基準となる音量に関する音声データとの比較の結果、抽出された音量が評価基準となる音量より大きいほど、キャラクタの振る舞いを大げさにし、抽出された音量が評価基準となる音量より小さいほど、キャラクタの振る舞いをひかえめにすることを特徴とするキャラクタの動作を制御する方法。

【請求項 1 5】

プレイヤからの音声入力を受け付ける音声入力装置を利用可能なエンタテインメント装置で実行できるプログラムを記録した記憶媒体であって、

前記プログラムは、

上記音声入力装置を介して受け付けたプレイヤの音声から、相対音程に関する情報を抽出する音程抽出処理と、

前記音声入力装置を介して受け付けたプレイヤの音声から、音量に関する情報を抽出する音量抽出処理とを、

前記抽出した相対音程に関する情報と音量に関する情報を評価し、その評価結果に応じた動作をキャラクタに行わせるキャラクタ制御処理とをエンタテインメント装置に行わせることを特徴とする記憶媒体。

【請求項 1 6】

請求項 1 5 記載の記憶媒体であって、

前記プログラムは、

プレイヤが入力すべき音声の内容を出力するガイド表示処理をさらにエンタテインメント装置に行わせることを特徴とする記憶媒体。

【請求項 1 7】

請求項 1 6 記載の記憶媒体であって、

前記プログラムは、



プレイヤーが入力すべき音声に対して、相対音程および音量についての評価基準となる音声データを記憶する基準音声データを参照する処理をエンタテインメント装置に行わせ、

前記キャラクタ制御処理は、前記抽出された相対音程に関する情報および音量に関する情報と、前記評価基準となる音声データとを周期的に比較し、その比較結果に基づいてキャラクタの動作内容を決定することを特徴とする記憶媒体。

【請求項 1 8】

請求項 1 6 記載の記憶媒体であって、

前記プログラムは、

プレイヤーが入力すべき音声の表現態様を出力する表現態様表示処理をさらにエンタテインメント装置に行わせることを特徴とする記憶媒体。

【請求項 1 9】

請求項 1 7 記載の記憶媒体であって、

前記キャラクタの動作は、あらかじめ用意された画像データを再生することにより表され、

前記キャラクタ制御処理は、前記プレイヤーが入力すべき音声の内容を出力するタイミングと、プレイヤーが音声の入力を開始するタイミングとの差に基づいて、前記画像データの再生速度を変化させることを特徴とする記憶媒体。

【請求項 2 0】

請求項 1 7 記載の記憶媒体であって、

前記キャラクタ制御処理は、前記抽出された相対音程に関する情報と、前記評価基準となる相対音程に関する音声データとを比較した結果、抽出された相対音程が評価基準となる相対音程より高いほど、キャラクタの表情を大げさにし、抽出された相対音程が評価基準となる相対音程より低いほど、キャラクタの表情をひかえめにすることを特徴とする記憶媒体。

【請求項 2 1】

請求項 1 7 記載の記憶媒体であって、

前記キャラクタ制御処理は、前記抽出された音量に関する情報と、前記評価基準となる音量に関する音声データとを比較した結果、抽出された音量が評価基準

となる音量より大きいほど、キャラクタの振る舞いを大げさにし、抽出された音量が評価基準となる音量より小さいほど、キャラクタの振る舞いをひかえめにすることを特徴とする記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、プレイヤーからの音声入力をキャラクタの動作に反映させるエンタテインメント装置、キャラクタの動作を制御する方法、およびそのような処理をエンタテインメント装置に実行させるためのプログラムを記録した記憶媒体に関する。

【0002】

【従来の技術】

エンタテインメント装置等で行われるゲームでは、プレイヤーがコントローラ、キーボード等の外部入力装置を用いて、操作対象であるプレイヤーキャラクタ等に指示を行う場合が多かったが、最近ではマイクロフォン等の音声入力装置を用いてプレイヤーが指示を行うゲームが現れている。

【0003】

このようなゲームでは、プレイヤーの入力音声に対して、例えば、音声スペクトラムの分析、標準パターンとのパターンマッチング等の音声認識技術を用いて、その内容を判断し、プレイヤーの入力音声に対応した行動をキャラクタにとらせることによりゲームを進行させていく。

【0004】

【発明が解決しようとする課題】

しかし、音声認識、特に、プレイヤーの言葉を解釈して、その内容をゲームに反映させることは装置への負担が大きく、処理に時間がかかるため、ゲームの円滑な進行のネックとなる場合がある。特に、ゲームに登場するキャラクタが、プレイヤーの音声にリアルタイムに反応するようなゲームに、音声入力を適用しようとする場合にこの問題は顕著になる。

【0005】

このため、音声入力を行うゲームは、主として、プレイヤーとキャラクターが対話を行い、プレイヤーの音声入力に対して、ワンテンポ遅れて受け答えする、あるいは行動を起こすような、リアルタイム性が重要視されないゲームに限られており、多様性に欠けているという問題が生じている。

## 【 0 0 0 6 】

本発明の目的は、音声入力により、リアルタイムにキャラクターが反応するゲームを提供することにある。

## 【 0 0 0 7 】

## 【課題を解決するための手段】

上記課題を解決するため、本発明は、プレイヤーからの音声入力を受け付ける音声入力装置を接続可能な、もしくは、備えるエンタテインメント装置であって、ゲームキャラクターの動作を制御するキャラクター制御手段と、上記音声入力装置を介して受け付けたプレイヤーの音声から、相対音程に関する情報を抽出する音程抽出手段と、前記音声入力装置を介して受け付けたプレイヤーの音声から、音量に関する情報を抽出する音量抽出手段とを備え、前記キャラクター制御手段は、前記抽出した相対音程に関する情報と音量に関する情報とに基づいて、キャラクターに動作を行わせることを特徴とするエンタテインメント装置を提供する。

## 【 0 0 0 8 】

このように、プレイヤー音声から音量と音程の情報を抜き出して処理を行うため、エンタテインメント装置に過大な負荷をかけることなく、円滑にゲームを進行させることができる。

## 【 0 0 0 9 】

また、このエンタテインメント装置は、プレイヤーが入力すべき音声の内容を出力するガイド表示手段をさらに備えることができる。

## 【 0 0 1 0 】

また、プレイヤーが入力すべき音声に対して、相対音程および音量についての評価基準となる音声データを記憶する基準音声データ記憶手段をさらに備え、前記キャラクター制御手段は、前記抽出された相対音程に関する情報および音量に関する情報と、前記評価基準となる音声データとを周期的に比較し、その比較結果に

基いてキャラクタの動作内容を決定するようにすることができる。

【0011】

さらに、前記キャラクタの動作は、あらかじめ用意された画像データを再生することにより表されるものであり、前記キャラクタ制御手段は、前記プレイヤーが入力すべき音声の内容を出力するタイミングと、プレイヤーが音声の入力を開始するタイミングとの差に基いて、前記画像データの再生速度を変化させることとすることができる。

【0012】

さらに、前記キャラクタ制御手段は、前記抽出された相対音程に関する情報と、前記評価基準となる相対音程に関する音声データとを比較した結果、抽出された相対音程が評価基準となる相対音程より高いほど、キャラクタの表情を大げさにし、抽出された相対音程が評価基準となる相対音程より低いほど、キャラクタの表情をひかえめにする一方、前記抽出された音量に関する情報と、前記評価基準となる音量に関する音声データとを比較した結果、抽出された音量が評価基準となる音量より大きいほど、キャラクタの振る舞いを大げさにし、抽出された音量が評価基準となる音量より小さいほど、キャラクタの振る舞いをひかえめにするにとすることができる。

【0013】

【発明の実施の形態】

本発明の実施の形態について、図面を参照して詳細に説明する。

【0014】

まず、本発明の実施形態である音声入力操作システムに係るエンタテインメント装置10のハードウェア構成について図8に示すブロック図を参照して説明する。

【0015】

本図において、エンタテインメント装置10は、メインCPU100と、グラフィックスプロセッサ(GP)110と、I/Oプロセッサ(IOP)120と、CD/DVD読み出し部130と、サウンド再生処理プロセッサ(SPU)140と、サウンドバッファ141と、OS-ROM150と、メインメモリ16

0と、IOPメモリ170とUSBインタフェース175とを備える。

【0016】

そして、メインCPU100とGP110とは、専用バス101によって接続されている。また、メインCPU100とIOP120とは、バス102によって接続されている。また、IOP120とCD/DVD読み出し部130とSPU140とOS-ROM150とは、バス103に接続されている。

【0017】

メインCPU100には、メインメモリ160が接続され、IOP120には、IOPメモリ170が接続されている。さらに、IOP120には、コントローラ(PAD)180とUSBインタフェース175とが接続される。

【0018】

メインCPU100は、OS-ROM150に格納されたプログラム、あるいは、CD/DVD-ROM等からメインメモリ160に転送されたプログラムを実行することによって、所定の処理を行う。

【0019】

GP110は、本エンタテインメント装置のレンダリング機能を受け持つ描画プロセッサであり、メインCPU100からの指示に従って、描画処理を行う。

【0020】

IOP120は、メインCPU100と、周辺装置、例えばCD/DVD読み出し部130、SPU140等との間のデータのやり取りを制御する入出力用サブ・プロセッサである。

【0021】

CD/DVD読み出し部130は、CD/DVDドライブに装着されたCD-ROM、DVD-ROMからデータを読み出し、メインメモリ160に設けられたバッファ領域161への転送を行う。

【0022】

SPU140は、サウンドバッファ141に格納された圧縮波形データ等を、メインCPU100等からの発音命令に基づいて、所定のサンプリング周波数で再生する。

## 【0023】

OS-ROM150は、起動時にメインCPU100およびIOP120が実行するプログラム等が格納されている不揮発性メモリである。

## 【0024】

メインメモリ160は、メインCPU100の主記憶装置であり、メインCPU100が実行する命令、および、メインCPU100が利用するデータ等が格納される。またメインメモリ160には、CD-ROM、DVD-ROM等の記録媒体から読み出されたデータを一時的に格納するバッファ領域161が設けられている。

## 【0025】

IOPメモリ170は、IOP120の主記憶装置であり、IOP120が実行する命令、および、メインCPU100が利用するデータ等が格納される。

## 【0026】

コントローラ(PAD)180は、操作者の指示を受付けるインタフェースである。

## 【0027】

USBインタフェース175には、USBマイクロフォン17が接続される。USBマイクロフォン17は、プレイヤーの音声が入力されると、所定のサンプリング周波数と量子化ビット数によるA/D変換等を施して、USBインタフェース175に音声データを送る。

## 【0028】

図9は、エンタテインメント装置10の使用状態を説明する図である。本図において、エンタテインメント装置本体11のコネクタ部12にコントローラ(PAD)180が接続される。また、エンタテインメント装置本体11の画像音声出力端子13には、画像音声出力用ケーブル14が接続される。このケーブル14の他端にはテレビ受信装置等の画像音声出力装置15が接続される。エンタテインメント装置の操作者は、コントローラ(PAD)180を用いて操作を指示する。エンタテインメント装置10は、コントローラ180を介して、操作者からの指示を受け付け、この指示に対応した画像データと音声データを、画像音声

出力装置 15 に出力する。そして、画像音声出力装置 15 が画像と音声を出力する。

#### 【0029】

また、エンタテインメント装置本体 11 の USB コネクタ 16 には、USB マイクロフォン 17 が接続され、プレイヤーの音声入力を受け付ける。

#### 【0030】

次に、本実施例の音声入力操作システムの構成について、図 1 のブロック図を参照して説明する。図示するように音声入力操作システムは、制御部 201 と入力制御部 202 と表示制御部 203 とシナリオデータ 301 と動画データ 302 と基準音声データ 303 とを備えて構成される。

#### 【0031】

制御部 201 は、ゲーム制御機能 2011 と、字幕制御機能 3012 と、入力音声評価機能 2013 と、動画制御機能 2014 とを有しており、メインメモリ 160 等に格納されたプログラムを、主としてメイン CPU 100 が実行することで、メイン CPU 100 上等に構築され、それぞれの機能を実現する。

#### 【0032】

ゲーム制御機能 2011 は、シナリオデータ 301 を読み込み、あらかじめ定められたストーリーに基いて、ゲームを進行させる処理を行う。

#### 【0033】

ここで、シナリオデータ 301 は、必要に応じて CD-ROM、DVD-ROM 等の記録媒体から読み出されるデータで、例えば、ストーリー展開、プレイヤーが入力すべきセリフに関する字幕データ、プレイヤーの入力に対するキャラクタの応答等に関するデータが記録されている。これらのデータは、インデックス等を付されて管理されており、このインデックスをキーに、ストーリー展開に合わせて表示、再生される。

#### 【0034】

字幕制御機能 3012 は、ストーリー展開上のシーンに関連付けられてシナリオデータ 301 に記録されている字幕を、表示制御部 203 を介して表示装置に表示させる処理を行う。この字幕はプレイヤーに対して、セリフの音声入力を促すガ

イドとしての役割を担う。そして、ある時刻において、プレイヤーが音声入力すべき文字については、その旨がわかるようにハイライト処理等を施して（カラオケの歌詞ガイド表示のように）、表示装置に表示させる。

## 【 0 0 3 5 】

入力音声評価機能 2 0 1 3 は、マイクロフォン等の音声入力装置を介して入力されたプレイヤーの音声データを、基準音声データ 3 0 3 に記録されている基準音声と比較して評価を行う。

## 【 0 0 3 6 】

具体的には、プレイヤーが入力した音声に対し、所定間隔毎、例えば 1 0 分の 1 秒毎に F F T 等（これは、ソフトウェア的に実現することが可能で、例えば、制御部 2 0 1 内に構築しておくことができる）を用いて基本周波数（音の高さ）を抽出するとともに、音量（音圧）を測定する。なお、音の高さを把握する要素は基本周波数の抽出に限らず、例えば、音声スペクトルの第 2 フォルマント等を抽出して把握するようにしてもよい。

## 【 0 0 3 7 】

なお、プレイヤーの入力する音声は、一フレーズ、すなわち、一続きに入力されるセリフを 1 単位として扱うものとする。この単位は、字幕において、ひとかたまりで表示するようにしておくことにより、プレイヤーが認識することができる。

## 【 0 0 3 8 】

図 2 は、「こんにちは」というセリフを音声入力した場合の、基本周波数と音量の変化の一例をグラフに表した図である。このとき、プレイヤーのセリフの長さが、2 秒であったとすると、測定ポイントは 2 0 個となり、基本周波数と音量は 2 0 個の時系列データとなる。ここで、基本周波数、音量ともに 0 ～ 1 0 0 までの値で表すものとし、基本周波数は、最初の測定ポイントを 5 0 として相対量に変換し、音量は絶対量として表すものとする。もちろん、これらの数値は厳密なものである必要はなく、基本周波数、音量の変化の度合いが把握できる程度でよい。

## 【 0 0 3 9 】

この図から明らかなように、本システムにおいては、プレイヤー音声の音量と音



程を把握するようにしており、発音は判断していない。例えば「こーんにちは」と「こんにちーは」を、同じ時間をかけて、時間的な音量変化、抑揚変化を同じくして入力した場合は、システムは同じ音声が入力されたものと把握する。さらには「あーーーー」と入力しても同様である。

## 【0040】

このように、本システムでは発音は評価しないため、音声入力を扱うにもかかわらず、処理に過度な負担をかけることなくゲームを実行することができる。すなわち、一般に、音程、音量の抽出は容易であり、ほぼリアルタイムに行うことができるため、ゲームの処理速度にもほとんど影響を与えることはない。

## 【0041】

基準音声データ303には、プレイヤーが入力したセリフの評価の基準となる音声データが記録されており、上記の所定期間毎にサンプリングした基本周波数の変化と、音量の変化がデータ化されている。

## 【0042】

入力音声評価機能2013は、プレイヤーが入力すべきセリフに対応する基準音声を基準音声データ303から抽出して、プレイヤーからの音声入力開始されるのを検知すると、セリフの開始時点を中心に所定期間毎にその差を算出して、評価を行う。

## 【0043】

例えば、セリフ「こんにちは」の基準音声データ303が図3に示すグラフの破線で表され、プレイヤーの入力データが実線で表されるものとする、その差は図4に示すグラフのようになる。

## 【0044】

そして、入力音声評価機能2013は、この差を基に、所定期間毎に入力音声の評価を行う。

## 【0045】

ここで、基本周波数の変化は、セリフの高低の変化、すなわち抑揚として把握し、プレイヤーの音声と基準音声との差に基づく評価をキャラクターの表情の変化に反映させるものとする。なお、基本周波数の変化を相対量としたのは、個人の声の

基本的な高さの差を考慮したためである。

【 0 0 4 6 】

音量の変化は、感情移入度として把握し、プレイヤーの音声と基準音声との差（テンション値）に基く評価をキャラクタの手足等の振る舞いの変化に反映させるものとする。

【 0 0 4 7 】

図4に示すグラフにおいては、プレイヤーの音声が基準音声より値が高いほど、ハイテンションで、プレイヤーの音声が基準音声より値が低いほどローテンションになるような評価基準としている。そして、プラスマイナスの度合いが大きいほど、テンションの度合いも大きくなるようになっている。

【 0 0 4 8 】

図4に示す例では、音量評価においては、セリフ内においてハイテンションとローテンションを行き来し、音程変化においては、全体的にハイテンションとなっている。

【 0 0 4 9 】

入力音声評価機能 2 0 1 3 は、上述のような所定期間毎に評価を行うとともに、一フレーズ毎の評価も行う。これは「こんにちは」のセリフ全体としてどの程度基準音声と乖離しているかを評価するもので、例えば、図4において、各所定期間毎の差の値について±0の線からの距離（各測定ポイントのプレイヤー音声と基準音声の差の絶対値）を求め、その総和とすることができる。このときは、値が少ないほうが、より基準音声に近いものとして評価が高いものとなる。

【 0 0 5 0 】

図5は、音量評価を例に、上述の入力音声評価機能 2 0 1 3 の評価についてまとめたものである。ここでは、簡単のため、5個の測定ポイントで評価を行うものとする。本図において、測定ポイント毎、すなわち所定期間毎のプレイヤー入力音声と基準音声の差であるテンション値は、+10、±0、-10、-20、+10と遷移し、フレーズの評価は距離（絶対値）の合計である50となる。

【 0 0 5 1 】

なお、プレイヤー入力音声と基準音声の終了時刻が同じでなく、どちらかが早く

終了してしまった場合は、例えば、所定期間毎の評価については、一方が終了したら評価を終了するようにして、フレーズの評価については、セリフのスピードが正確でないとして、一方が終了した場合、その後の差を最大値であると仮定して、評価を悪くするようにすることができる。

## 【0052】

動画制御機能2014は、キャラクターの動作を記録した動画データ302を読み込み、入力音声評価機能2013の評価結果を、キャラクターの動作に反映させる処理を行う。

## 【0053】

ここで、動画データ302は、必要に応じてCD-ROM、DVD-ROM等の記録媒体から読み出されるデータで、ストーリー展開に対応した、キャラクターの動作に関するが記録されている。動画データ302に記録されたキャラクター、特にプレイヤーの操作対象のキャラクターの動作に関するデータでは、表情、感情等を表現するための動き、例えば、目の大きさ、口の開き具合、身振りの大きさ等をパラメータで変更できるようになっている。

## 【0054】

例えば、動画データ302として、キャラクターが驚くという動作に関するデータが記録されているときに、目の大きさ、口の開き具合、手の動きに関するパラメータが変更できるものとする。このときのパラメータの内容として、「大げさ」「普通」「ひかえめ」であるとする。このときのパラメータの変更に伴うキャラクターの動作の変化の例を図6に示す。

## 【0055】

本図において、(a)は、目の大きさ、口の開き具合、手の動きともに「大げさ」とした場合の動作である。(b)は、目の大きさ、口の開き具合、手の動きともに「普通」とした場合の動作である。(c)は、目の大きさ、口の開き具合、手の動きともに「ひかえめ」とした場合の動作である。(d)は、目の大きさ、口の開き具合は「普通」、手の動きは「大げさ」とした場合の動作である。

## 【0056】

このように、動画制御機能2014は、キャラクターの動作をパラメータの組合

せによって変化させることができる。

【0057】

入力制御部202は、入力装置として接続されたマイクロフォンからの入力音声信号の制御等を行う。

【0058】

表示制御部203は、メインCPU100等の指示にしたがってGP110上に構築され、制御部201から受け付けた画像データをゲーム処理部802から渡された画面データに基いて表示画面データを生成する。生成された表示画面データは表示装置に出力され、これを受けた表示装置は、表示画面上に、表示画面データにしたがった画像を表示する。

【0059】

次に、本実施例におけるエンタテインメント装置10の動作について説明する。

【0060】

ゲームが開始されると制御部201は、シナリオデータ301を読み込み、シナリオに関連付けられた動画データ302を再生する。そして、プレイヤーが担当するキャラクターがセリフを言うシーンになると、そのセリフの字幕を表示して、プレイヤーに音声を入力を促す。

【0061】

図7はこのときの処理の流れを説明するフロー図である。

【0062】

まず、前述のように制御部201は、プレイヤーが入力すべきセリフを表示制御部203を介して表示装置に表示させる（S101）。

【0063】

そして、制御部201は、この字幕に対して、読むべき文字をハイライト表示してプレイヤーに音声入力を促す（S102）。

【0064】

プレイヤーは、このハイライト表示に合わせてセリフを入力するものとする。

【0065】

なお、この字幕にあわせて、セリフの入力に関する情報を表示させるようにしてもよい。例えば、そのシーンが衝撃的な場面で、基準音声が大げさに記録されているとすると、プレイヤーが大げさなセリフ入力を課すために、「おおげさに」という表現のガイドを表示するようにする。このとき、プレイヤーは表現のガイドどおりにセリフを入力することで、高い評価を得られるようになる。

## 【 0 0 6 6 】

制御部 2 0 1 は、字幕の最初の文字をハイライト表示する時点の前後の所定時間、例えば 1 秒、あるいは、2 秒以内の範囲に、プレイヤーがセリフ入力を開始した場合に、そのセリフに対する入力を有効なものとして扱うものとする（S 1 0 3）。この範囲より先にセリフ入力を開始した場合、あるいは、この範囲より後にセリフ入力を開始した場合は、無効であるとして扱い、その入力に対する評価を下げるものとする（S 1 0 4）。

## 【 0 0 6 7 】

また、制御部 2 0 1 は、前記の有効期間内にセリフ入力開始された場合に、字幕の最初の文字をハイライト表示する時点より、セリフ入力開始時点が早いときは、そのセリフ入力に係るシナリオに関連付けられた動画データ 3 0 2 の再生速度を遅くして再生する（S 1 0 6）。一方、字幕の最初の文字をハイライト表示する時点より、セリフ入力開始時点が遅いときは、そのセリフ入力に係るシナリオに関連付けられた動画データ 3 0 2 の再生速度を早くして再生する（S 1 0 7）。再生速度を速める度合い、遅める度合いは、字幕の最初の文字をハイライト表示する時点とセリフ入力開始時点との差に比例する。

## 【 0 0 6 8 】

すなわち、制御部 2 0 1 は、プレイヤーのセリフ入力開始時点が、字幕の開始時点とずれた場合であっても、そのセリフが終了するタイミングと、そのセリフに対するキャラクタの動作が終了するタイミングをそろえるように調整する。

## 【 0 0 6 9 】

制御部 2 0 1 は、プレイヤーの音声入力に対して、所定期間毎、例えば 1 0 分の 1 秒毎に前述の評価を行う（S 1 0 8）。そして、その評価に基づいて、キャラクタの動作を即座に調整して映像に反映させる（S 1 0 9）。この動作を、プレイ

ヤの音声入力終了するまで繰り返す（S 1 1 0）。

【 0 0 7 0 】

このときの処理について説明する。

【 0 0 7 1 】

制御部 2 0 1 は、前述のように所定期間毎に、プレイヤー音声と基準音声の音量と音程の差を算出する。そして、この差の値をそれぞれ音量テンション値、音程テンション値とする。

【 0 0 7 2 】

動画データ 3 0 2 には、テンション値とキャラクタの動作のパラメータが関連付けられている。例えば、キャラクタが驚くという動作のときは、音量テンション値はキャラクタの手の動きに関連付けられ、音量テンション値が - 2 5 未満の場合は「ひかえめ」で、音量テンション値が - 2 5 以上 + 2 5 未満の場合は「普通」で、音量テンション値が + 2 5 以上の場合は「大げさ」とする。

【 0 0 7 3 】

また、音程テンション値は、キャラクタの目の大きさと口の開き具合に関連付けられ、音程テンション値が - 2 5 未満の場合は「ひかえめ」で、音程テンション値が - 2 5 以上 + 2 5 未満の場合は「普通」で、音程テンション値が + 2 5 以上の場合は「大げさ」とする。

【 0 0 7 4 】

制御部 2 0 1 は、所定期間毎に音量テンション値と音程テンション値とを算出し、その値に基いて、動作のパラメータ内容を決定する。例えば、音量テンションが + 3 0 で、音程テンションが + 1 0 であったとすると、手の動きは「大げさ」で、目の大きさと、口の開き具合は「普通」となる。

【 0 0 7 5 】

そして、制御部 2 0 1 は、このパラメータ内容に応じた画像を生成して、表示制御部 2 0 3 を介して表示装置に表示させる。

【 0 0 7 6 】

この処理を、プレイヤーのセリフ入力終了するまで、所定期間毎に行うことにより、プレイヤーの入力にリアルタイムに反応した動作をキャラクタに行わせるこ

とができる。なお、プレイヤーが入力すべきセリフは一フレーズを単位としているため、音声入力の入力開始を検知した後、所定期間、例えば0.5秒間プレイヤーからの音声が無かった場合にセリフ入力終了したものと判断するようにすることができる。

## 【0077】

制御部103は、プレイヤーのセリフ入力終了すると、前述のように、そのセリフ全体に対する評価を行う（S111）。これは、基本音声にどれだけ近い音声を入力することができたかを示す評価である。この評価を、例えば、あるストーリーを通じて累積していき、その結果、一定の評価を得ることができなかった場合には、次のストーリーに進めないようにして、ゲーム性を高めるようにすることができる。

## 【0078】

なお、本発明は、上記の実施形態に限定されるものではなく、その要旨の範囲内で様々な変形が可能である。

## 【0079】

例えば、上記の例では、テンション値をパラメータの内容（「大げさ」「普通」「ひかえめ」）と関連付けて、キャラクターの行動を定めるようにしていたが、テンション値と、キャラクターの行動とを直接関連付けて（例えば、目の大きさを0～100に区分し）、テンション値をそのままパラメータとして用いるようにしてもよい。

## 【0080】

また、エンタテインメント装置10の外観およびハードウェア構成は、図8、図9に示したものに限定されない。エンタテインメント装置10は、たとえば、CPUと、メモリと、ハードディスク装置などの外部記憶装置と、CD-ROM DVD-ROMなどの可搬性を有する記憶媒体からデータを読み取る読取装置と、キーボード、マウス、マイクロフォンなどの入力装置と、ディスプレイなどの表示装置と、インターネットなどのネットワークを介して通信を行うためのデータ通信装置と、上述した各装置間のデータ送受を司るインターフェースといった、一般的な電子計算機の構成を有するものであってもよい。この場合、エンタテ

インメント装置 10 上に、図 1 に示す構成を構築するためのプログラムおよび各種データは、読取装置を介して、可搬性を有する記憶媒体から読み出され、メモリや外部記憶装置に記憶されるようにしてもよいし、あるいは、データ通信装置を介して、ネットワークからダウンロードされ、メモリ、外部記憶装置に記憶されるようにしてもよい。

【0081】

【発明の効果】

上述のように、本発明によれば、音声入力により、リアルタイムにキャラクタが反応するゲームを実現することができる。

【図面の簡単な説明】

【図 1】は、本実施例の音声入力操作システムの構成を説明するブロック図である。

【図 2】は、セリフを音声入力した場合の、音程と音量の変化の一例をグラフに表した図である。

【図 3】は、音程と音量の変化について、プレイヤー音声と基準音声との関係の一例をグラフに表した図である。

【図 4】は、音程と音量の変化について、プレイヤー音声と基準音声の差をグラフに表した図である。

【図 5】は、音量評価を例に、入力音声評価機能 2013 の評価についてまとめた図である。

【図 6】は、パラメータの変更に伴うキャラクタの動作の変化の例を示す図である。

【図 7】は、プレイヤーからのセリフを受け付けるときの処理の流れを説明するフロー図である。

【図 8】は、エンタテインメント装置 10 のハードウェア構成について説明するブロック図である。

【図 9】は、エンタテインメント装置 10 の使用状態を説明する図である。

【符号の説明】

101…制御部、102…入力制御部、103…表示制御部



2 0 1 …シナリオデータ、2 0 2 …動画データ、2 0 3 …基準音声データ

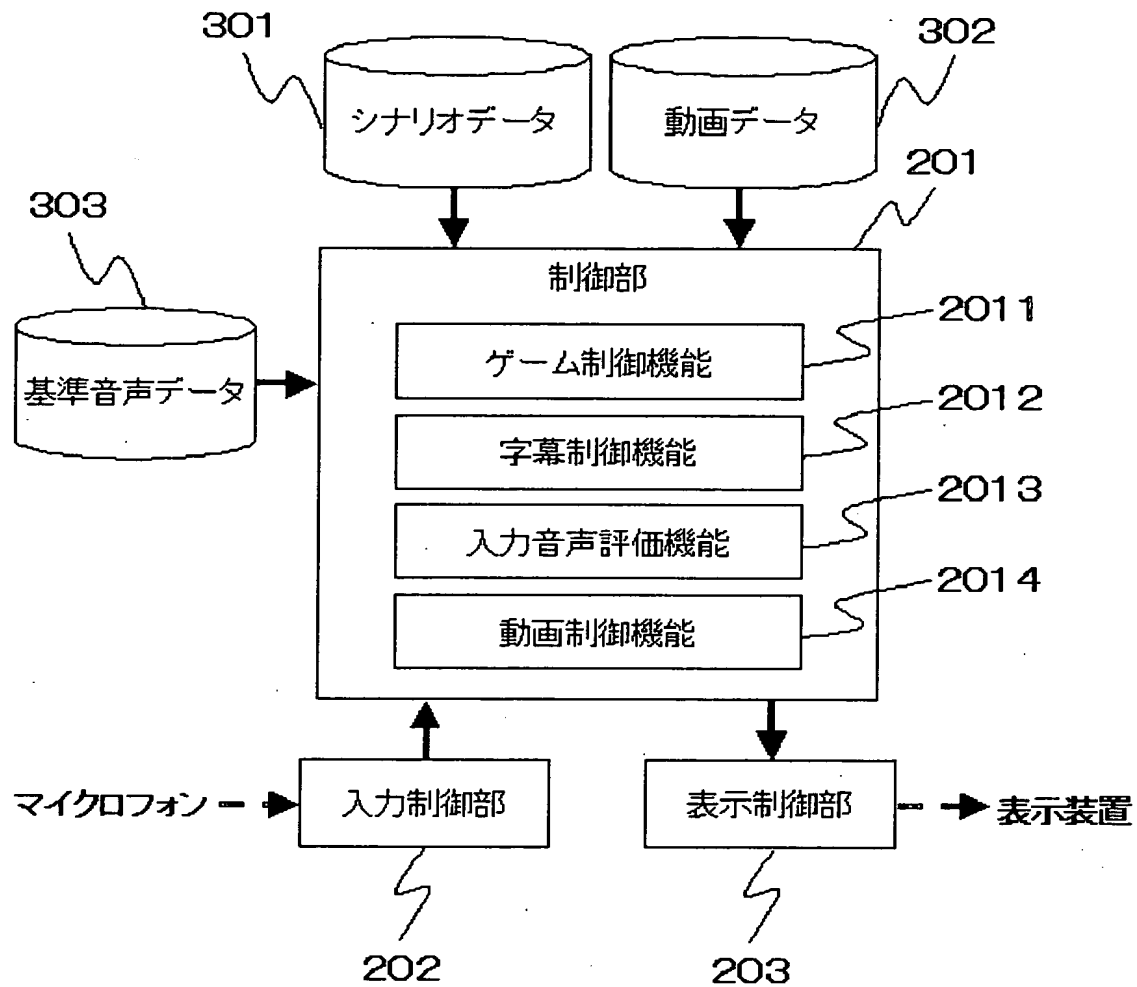
1 0 1 1 …ゲーム制御機能、1 0 1 2 …字幕制御機能

1 0 1 3 …入力音声評価機能、1 0 1 4 …動画制御機能

【書類名】図面

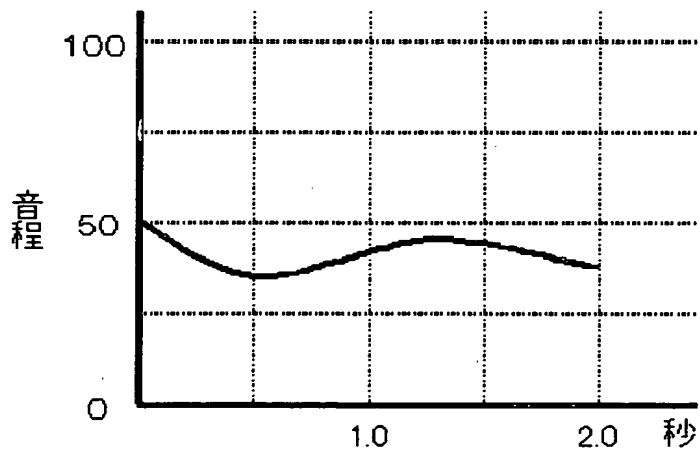
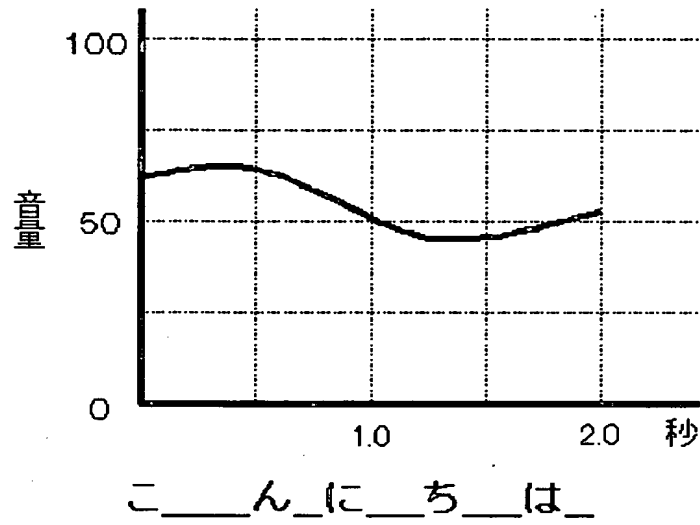
【図1】

図1



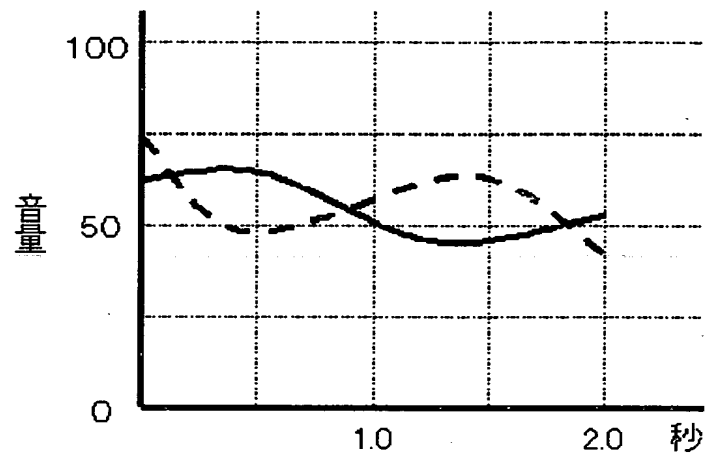
【図2】

図2

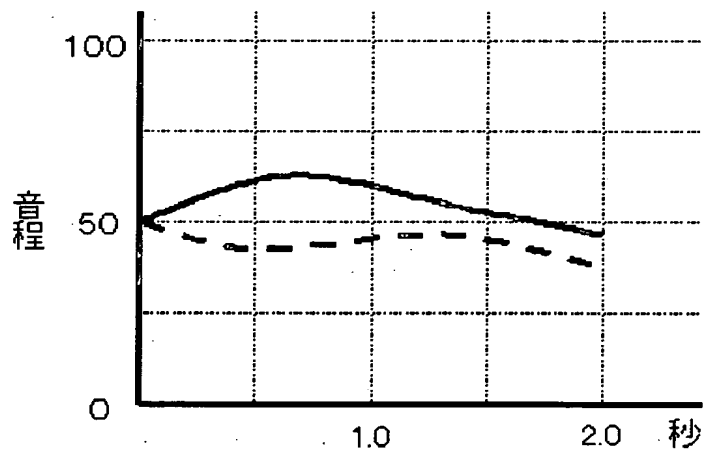


【図3】

図3

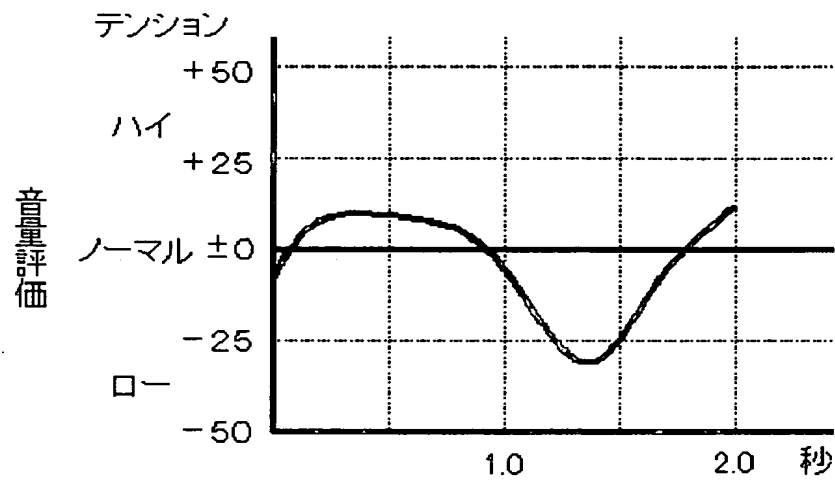


こ\_\_ん\_\_に\_\_ち\_\_は\_\_

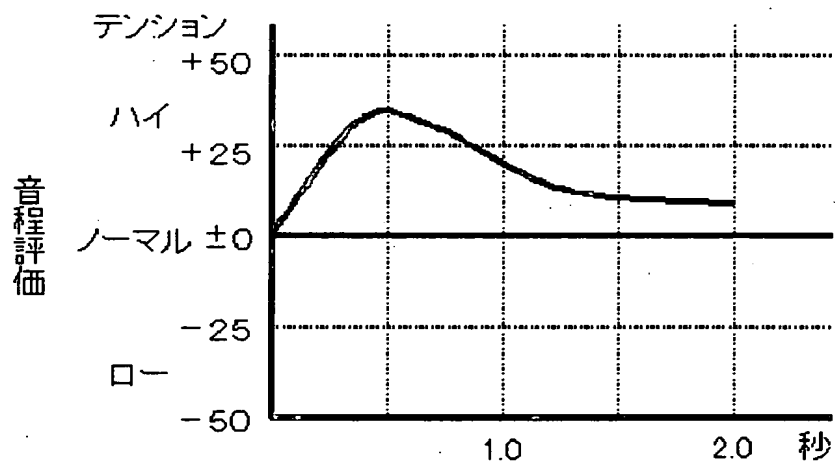


【図4】

図4



こ\_\_ん\_\_に\_\_ち\_\_は\_\_



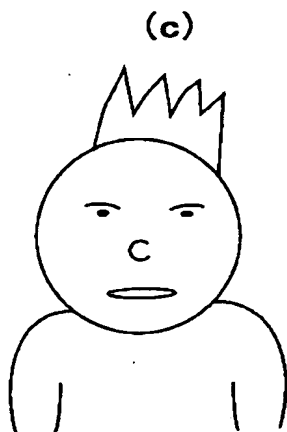
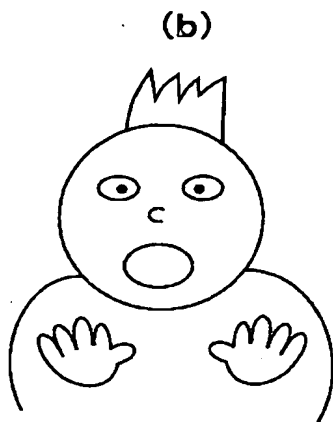
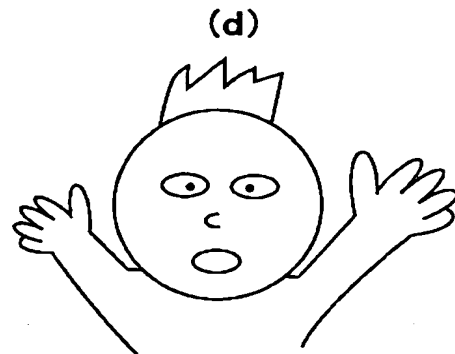
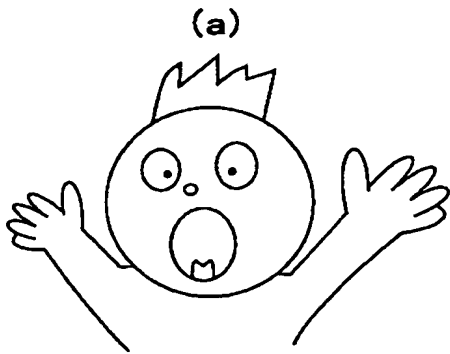
【図 5】

図5

プレイヤー入力	基準音声	差	距離
50	40	+10	10
40	40	±0	0
50	60	-10	10
60	80	-20	20
70	60	+10	10
			合計 50

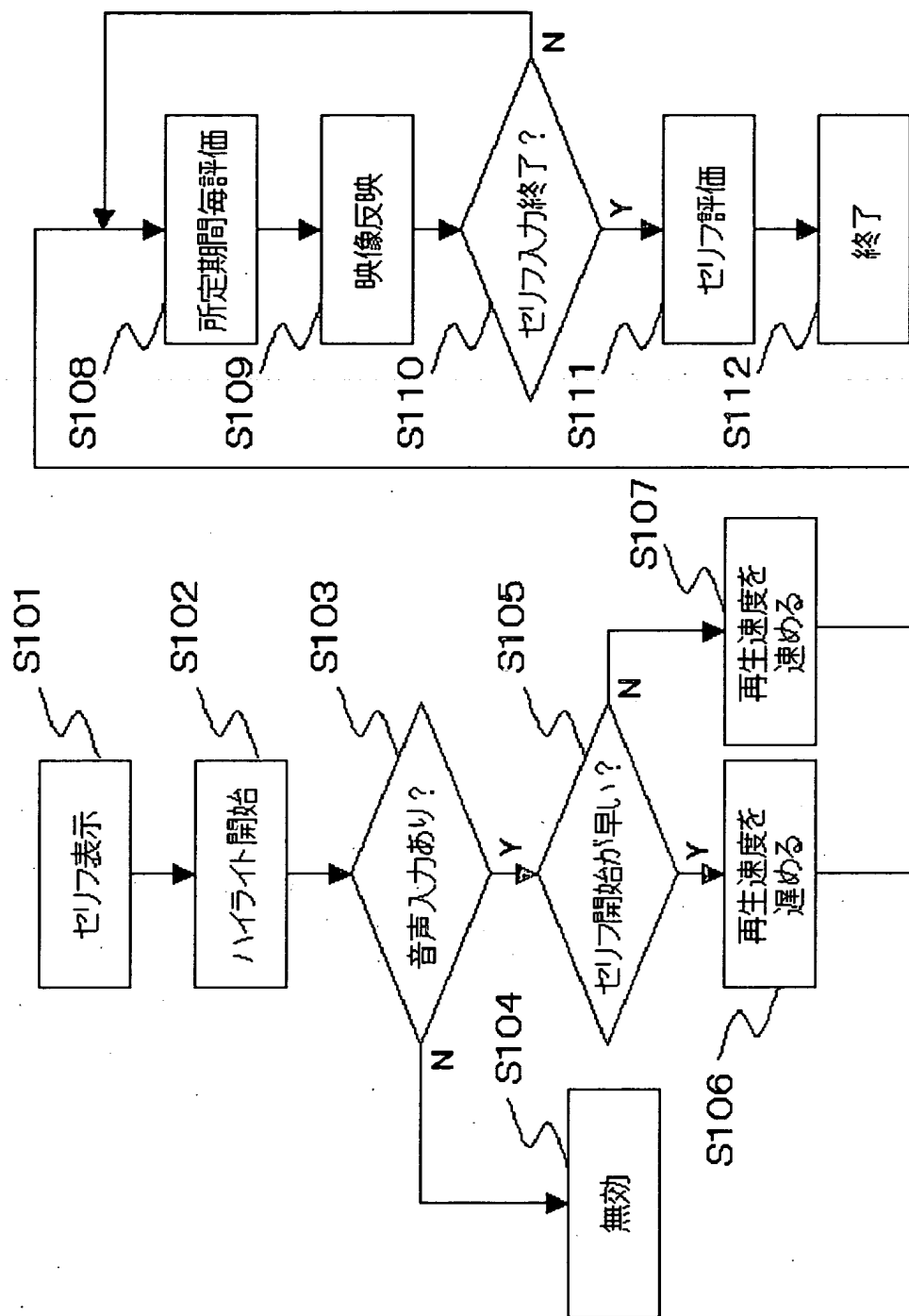
【図 6】

図 6



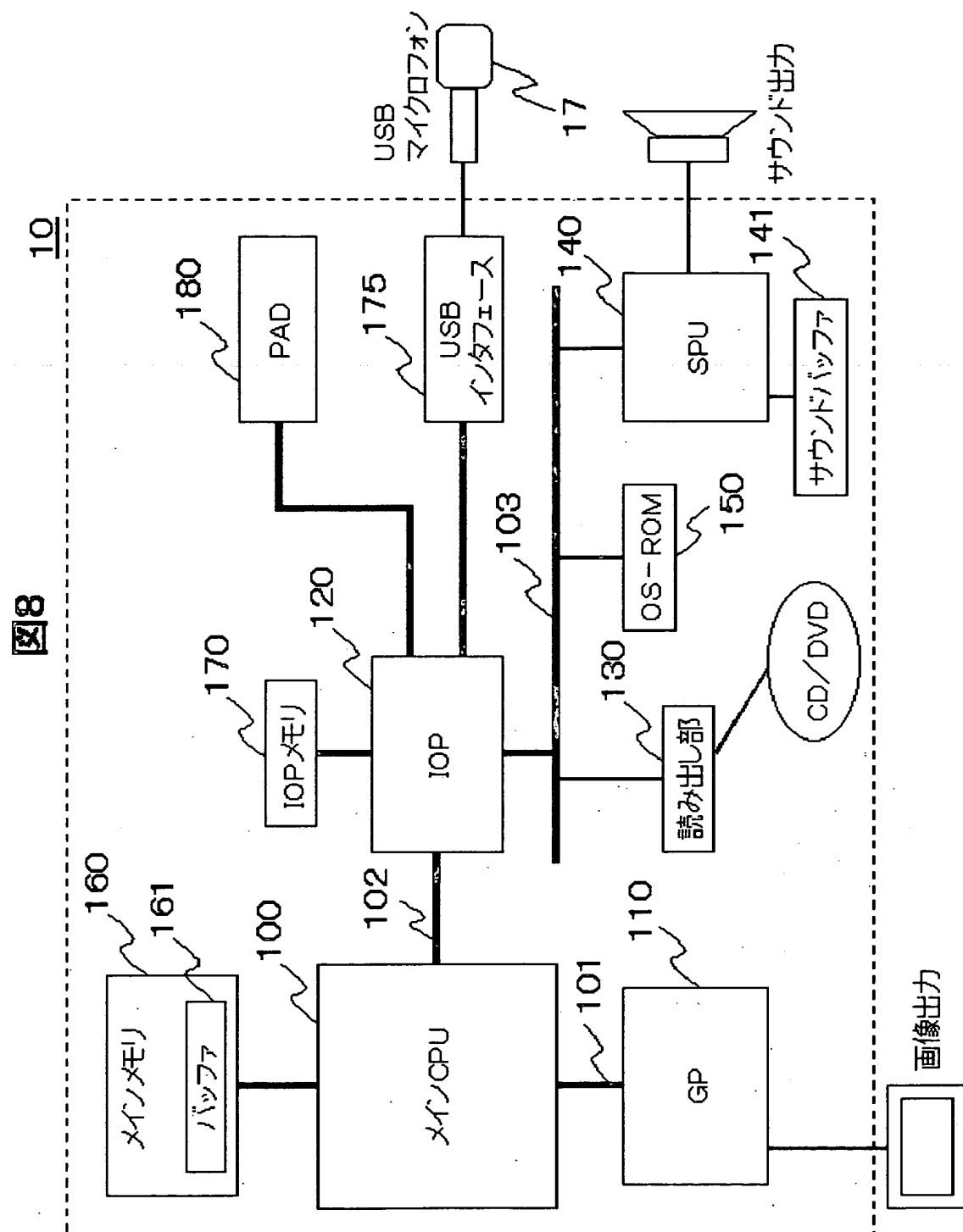
【図7】

図7



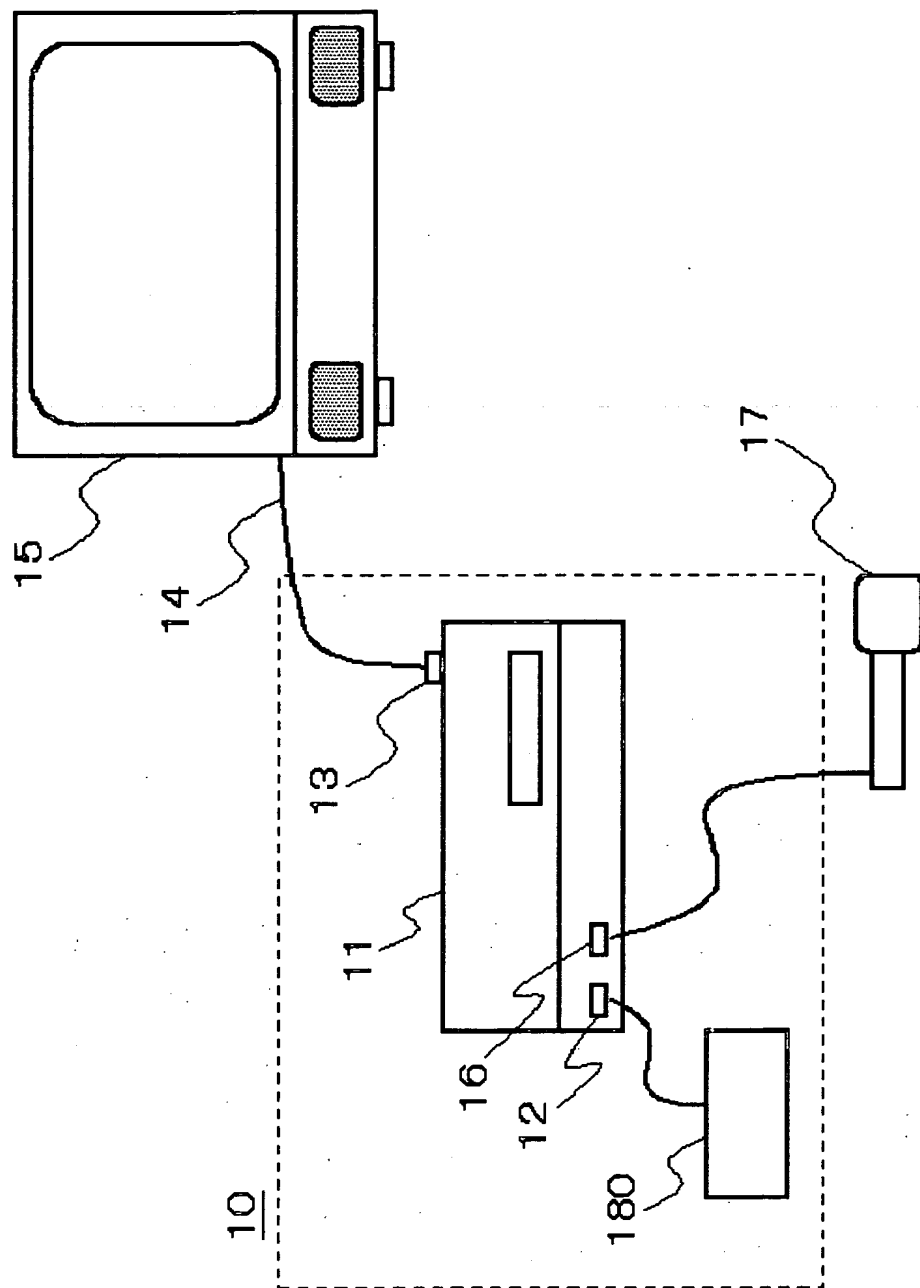


【図8】



【図 9】

図 9



【書類名】 要約書

【要約】

【課題】 プレイヤからの音声入力により、リアルタイムにキャラクターが反応するゲームを実現する。

【解決手段】 マイクロフォンを介して入力されたプレイヤの音声から、音程と音量を抽出して、セリフにおける音程と音量の変化を把握する。このデータと基準音声データ 2 0 3 に記録されている基準データとの差を求め、その差に基づいて入力されたセリフの評価を行う。プレイヤの操作対象となるキャラクターは、その評価により、動作の内容が定められ、リアルタイムに反応する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [395015319]

1. 変更年月日 1997年 3月31日

[変更理由] 住所変更

住 所 東京都港区赤坂7-1-1

氏 名 株式会社ソニー・コンピュータエンタテインメント